Amendments to the Claims

This listing of claims will replace all prior versions, and listings, of claims in the application:

Claim 1 (withdrawn) A computer implemented method for determining the direction of a consensus sequence of a cluster of sequences with contradictions in directions comprising:

determining the probability (b) that the contradictions are explained by random errors according to a statistical model and the weighted number of contradictory sequences in the cluster; and

defining the direction of majority of the sequences as the direction of the consensus sequence if the probability is the same as or greater than a threshold value (T) and x≠n/2.

Claim 2 (withdrawn) The method of Claim 1 wherein the statistical model is a binomial distribution and the probability is calculated as follows:

$$b(x;n,p) = \frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}$$

wherein n is the weighted number of the sequences in the cluster; p is the probability of random errors resulting in the contradictions; and x is the number of the contradictory sequences.

Claim 3 (withdrawn) The method of Claim 2 wherein CDS and mRNA sequences carry a higher weight than 5' EST or 3' EST; directionless EST carrys a weight of 0.

Claim 4 (withdrawn) The method of Claim 2 wherein the weights to different types of sequences are the same.

Claim 5 (withdrawn) The method of Claim 2 wherein the threshold value is around 0.001.

Claim 6 (withdrawn) The method of Claim 2 wherein the threshold value is around 0.002.

Claim 7 (withdrawn) The method of Claim 2 wherein the threshold value is around 0.003.

Claim 8 (withdrawn) The method of Claim 2 further comprising defining the direction of majority of the sequences as the direction of the consensus sequence if the probability is lower than the threshold value and $x \leq n*(P_t)$.

Claim 9 (withdrawn) The method of Claim 8 further comprising further subclustering for the minority direction and majority direction if the probability is smaller than the threshold value and $x > n*P$.

Claim 10 (withdrawn) The method of Claim 9 wherein the p is between 0.03-0.10.

Claim 11 (withdrawn) The method of Claim 10 wherein the p is around 0.06.

Claim 12 (withdrawn) The method of Claim 11 wherein the p is determined according to binomial frequency distribution of contradictory sequences in a plurality of clusters or subclusters of sequences.

Claim 13 (currently amended): A method of ~~selecting~~ obtaining candidate sequences for designing a probe array comprising:
cleaning raw sequences;
refining clusters of the raw sequences; ~~and~~
generating candidate design sequences, wherein the candidate design sequences are exemplar or consensus sequences of the clusters[[.]] ; and
selecting probes targeting the exemplar or consensus sequences for designing the probe array.

Claim 14 (currently amended): The method of Claim 13 wherein the cleaning comprises removing withdrawn sequences;
screening and filtering and masking raw sequences; and

~~triming~~ <u>trimming</u> terminal ambiguous sequence regions.

Claim 15 (original): The method of Claim 13 wherein the refining includes two level clustering.

Claim 16 (original): The method of Claim 13 wherein the generating comprises:
  selecting exemplary sequences.

Claim 17 (original): The method of Claim 16 wherein the generating comprises:
  generating alignments of sequences within clusters;
  calling consensus sequence bases according to consensus calling rules; and
  determining consensus sequence direction.

Claim 18 (original): The method of Claim 17 wherein the determining comprises defining the direction of sequences in the clusters as the consensus sequence direction if there is no contradictory sequence directions.

Claim 19 (currently amended): The method of Claim 18 wherein the determining further comprises
  determining the probability (b) that the contradictions are explained by random errors according to a statistical model and the weighted number of contradictory sequences in the cluster; and
  defining the direction of majority of the sequences as the direction of the consensus sequence if the probability is the same as or greater than a threshold value (T) and x≠n/2<u>, wherein x is the number of the contradictory sequences and n is the weighted number of the sequences in the cluster</u>.

Claim 20 (original): The method of Claim 19 wherein the statistical model is a binomial distribution and the probability is calculated as follows:

$$b(x;n,p) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

wherein n is the weighted number of the sequences in the cluster; p is the probability of random errors resulting in the contradictions; and x is the number of the contradictory sequences.

Claim 21 (currently amended): The method of Claim 20 wherein <u>coding regions (CDS)</u> and mRNA sequences carry a higher weight than 5' EST or 3' EST; directionless EST ~~carrys~~ <u>carries</u> a weight of 0.

Claim 22 (original): The method of Claim 21 wherein the weights to different types of sequences are the same.

Claim 23 (original): The method of Claim 22 wherein the threshold value is around 0.001.

Claim 24 (original): The method of Claim 22 wherein the threshold value is around 0.002.

Claim 25 (original): The method of Claim 22 wherein the threshold value is around 0.003.

Claim 26 (currently amended): The method of Claim 22 further comprising defining the direction of majority of the sequences as the direction of the consensus sequence if the probability is lower than the threshold value and $x \leq n^*(P_t)$, <u>wherein P_t is a threshold value of x/n above which non-random errors dominate over random errors</u>.

Claim 27 (currently amended): The method of Claim 26 further comprising further subclustering for the minority direction and majority direction if the probability is smaller than the threshold value and $x > n^* \text{~~P~~} \underline{P_t}$.

Claim 28 (original): The method of Claim 27 wherein the p is between 0.03-0.10.

Claim 29 (original): The method of Claim 27 wherein the p is around 0.06.

Claim 30 (original): The method of Claim 27 wherein the p is determined according to binomial frequency distribution of contradictory sequences in a plurality of clusters or subclusters of sequences.

Amendments to the Drawings

Applicants submit one sheet of formal drawings, containing amended Figure 7.

Currently amended Figure 7 incorporates the corrections made to original Figure 7 as

shown in the marked-up copy of the same filed on 10/03/02, which were approved by the

Examiner.

Attachment: Replacement Sheet (1 sheet)